REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-06-2009	Final Technical Report	23-09-2008 - 22-06-2009

4. TITLE AND SUBTITLE
Stochastic Model-Based Control of Multi-Robot Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W911NF-08-1-0503

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dejan Milutinovic and Devendra P. Garg

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Robotics and Manufacturing Automation (RAMA) Laboratory
Department of Mechanical Engineering and Materials Science
Duke University, Box 90300
Durham, NC, 27708

8. PERFORMING ORGANIZATION REPORT NUMBER
DU-RAMA-ARO-103

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 1211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See (*) in the field 14. ABSTRACT

14. ABSTRACT
In this report we consider control of single- and multi-robot systems as an optimal control problem. Solution of this problem may be of enormous complexity because of a large-number of robots, a large number of redundant states, and environment uncertainties. Motivated by estimation methods based on statistical sampling employed for solving complex estimation problems, we explore the possibility of using stochastic process samples for computing the optimal control. This approach can ultimately provide small-size, low-cost and efficient computational hardware for solving complex multi-robot control problems and in which computations are driven by laws of statistical physics.

(*) The views, opinions and/or findings contained in this report are those of the authors and should not be considered as an official Department of Army position, policy or decision, unless so designated by other documentation.

15. SUBJECT TERMS
multi-robot systems; stochastic optimal control; statistical sampling and estimation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
				35	

Preface

This report describes the research carried out during six month of 2008-2009 in the Robotics and Manufacturing Automation (RAMA) Laboratory of Duke University. The major effort during this time was the development of a theoretical approach for using samples of stochastic processes for solving the stochastic optimal control problems. This effort is described in the following sections.

During this time period a technical paper dealing with force field estimation based on video data was also prepared and submitted for publication consideration in a Special Issue on Physical System Modeling of the Journal of Dynamic Systems, Measurement, and Control, Series G of the ASME Transactions. The submitted paper's abstract is attached in Appendix A. Publication decision on this paper is still awaited. A scale-down copy of the poster presented at the CMPI Symposium on Multi-Scale Modeling of Host/Pathogen Interactions, June 23 - 25, 2009, Pittsburgh, PA, is attached as Appendix B.

The financial support provided by the United States Army Research Office for the Short Term Innovative Research (STIR) project entitled "Stochastic Model-Based Control of Multi Robot Systems" under Grant Number W911NF-08-1-0503 is gratefully acknowledged.

Abstract

In this report, we consider control of single- and multi-robot systems as an optimal control problem. Solution of this problem may be of enormous complexity because of a large-number of robots, a large number of redundant states, and environmental uncertainties. Motivated by estimation methods based on statistical sampling and employed for solving complex estimation problems, we explore the possibility of using stochastic process samples for computing the optimal control.

In our work, the state of an individual robot is described by the state composed of discrete and continuous state variables. We model the robot using a hybrid automaton and corresponding system of partial differential equations describing the robot's state probability density function. For solving the optimal control problem, we use the minimum-principle for partial differential equations to obtain the Hamiltonian that characterizes the optimal control. However, having in mind that the Hamiltonian evaluation based on the partial differential equations is computationally expensive, we propose and explore strategies in which the Hamiltonian evaluation is based on samples from stochastic processes generated from the individual robot model. We also show that if the analysis is limited to the so-called process self-adjoint dynamical systems, the Hamiltonian evaluation can be simplified.

Using computational statistical methods opens the possibility to solve control problems in robotics in real-time within the stochastic optimal control framework. This possibility also depends on processor design that implements computational methods. While the design of such processors is a separate issue, ideally it should be driven by stochastic processes, such as noises in electronic components. This concept can ultimately provide miniature computational hardware for solving complex multi-robot control problems, in which computations are driven by laws of statistical physics.

Contents

Nomenclature

$\mathbb{R}, \mathbb{R}^n$ - set of real numbers, set of real number vectors of dimension n.

Q - set of discrete states, i.e., integer indexes $\{1, 2, 3, \ldots\}$

U_{ad} - set of admissible control.

$\rho(x,t)$ - probability density function (PDF) of the hybrid state at time t. This variable is a vector of functions, it depends on $x \in X$ and $t \in R$, but x and t are frequently omitted in expressions.

$\rho_i(x,t)$ - the PDF component corresponding to the discrete state i, $i \in Q$.

$\pi(x,t)$ - the adjoint state distribution

$\bar{\pi}$ - the discrete approximation of the adjoint state distribution.

$\phi(x,t)$ - the adjoint state PDF

$P_i(x,t)$ - the probability of the discrete state i, $i \in Q$.

$P_i^{\pi}(x,t)$ - the probability of the discrete adjoint state i, $i \in Q$.

F_t - the transition rate matrix.

λ_{ij} - the component of the transition rate matrix $[F_t]_{ij} = \lambda_{ij}$.

F_u - the transition rate matrix that depends on control vector u, $u \in U_{ad}$.

F_∂ - the component of the linear operator F corresponding to the vector fields f_i of discrete states $i \in Q$.

$H(\rho, u, t)$ - the PDF, the control and time dependent Hamiltonian, ρ, u and t are frequently omitted.

u^* - the optimal control

E_ρ - the expected value with respect to the state PDF ρ.

List of Figures

Stochastic Model-Based Control of Multi-Robot Systems

1 Introduction

A multi-robot system is a collection of robots designed to perform a specified task. Collection of all relevant variables uniquely specifying the individual robot state, such as position, status, actuator powers etc., we call the state. In this respect, a solution of a multi-robot control problem assumes that the update of each individual robot state is in accordance with the specified multi-robot system task. The update is defined by the specific control approach or architecture. In this research, we have developed an approach in which the task specification is provided in the form of a cost function that should be maximized (or minimized) by the multi-robot system. Such cost functions can relate to the presence of robots at specific locations, or to the amount of information that can be gathered. It may also include other factors such as energy consumption or amount of communicated information.

From this perspective, the multi-robot system control can be considered as an optimal control problem resulting in the update of individual robot states leading towards a minimum (or maximum) of the cost function defining the multi-robot mission task. Solution of this problem may be of enormous complexity because of a large number of robots, a large number of redundant states, as well as existing environment uncertainties. In this report, we discuss an approach to deal with this complexity computationally for both single- and multiple-robot systems.

For many years it has been known that the optimal control and optimal estimation problems are dual [6]. For example, we use the optimal control theory to derive linear quadratic regulator (LQR), and in the same theoretical framework we can derive the Kalman filter (KF). In the first case we design the controller that minimizes the quadratic cost function over the state space, and in the second case we design the estimator minimizing the cost function involving state space estimations. Having this in mind, it is expected that estimation methods based on statistical sampling, employed for solving complex estimation problems [12], can be also exploited for solving complex control problems for single and more important multi-robot systems under presence of uncertainty.

Along this idea, Kappen et. al. [7, 13] used stochastic differential equations to model individual agents. Based on this description, it is possible to relate Hamilton-Jacobi-Bellman partial differential equation with samples of the stochastic process trajectories and use the samples to define the stochastic optimal control of multi-agent systems. In this framework, the state of individual robots is continuous. However, the state of real robots is generally described by a combination of continuous and discrete variables, i.e., by a hybrid state. Therefore, it is more natural to describe the robot behavior using a hybrid automaton [2]. The automaton describes the discrete and continuous variables change in time, which depends on events influencing the robot behavior. In the case of a large-size multi-robot population, it becomes highly complex to predict events from the robot local environment. Because of that, we model a large-size robots' population considering the stochastic hybrid model and study how it can be controlled.

In this research project, we consider a problem in which the presence of a large-size robot population, in a desired region of operating space, is maximized. This problem is formulated in the hybrid system framework in [11]. Its solution, based on the minimum principle for partial differential equations, is presented in [9, 10], but it is solved numerically only in the case when the presence of the robots is maximized along one dimension (1D). The difficulty in developing more realistic 2D control is rooted in the fact that that the Hamiltonian, which defines the optimal control, includes integral terms that depend on the solution of a system of partial differential equations (PDE). This system of PDEs is in general difficult to evaluate. To overcome this difficulty, we recognize that the PDE system solution contributing to the Hamiltonian is related to the stochastic process that can be computationally generated. Consequently, we propose to use samples of the stochastic processes to evaluate the Hamiltonian components. In this way, our approach is considerably different from stochastic optimal control work presented in [8]. There the stochastic processes have been used only as an analytical tool to map stochastic process to be controlled into the finite state space, in which the optimization is performed.

The benefit of using solution based on sampling, i.e, computational statistical methods, is that the control problems in robotics could be solved efficiently (in real-time) in an optimal control framework. This possibility also depends on the ability to implement sampling and computations with samples into the processor computing the control. While the design of such processors is a separate issue, ideally it should be driven by real stochastic processes such as a noise in electronic components. This concept can ultimately provide miniature computational hardware for solving complex multi-robot control problems. In this respect, the ultimate processor will be a system in which computations are based on laws of statistical physics.

In Section 2 of this report we discuss hybrid systems modeling and stochastic optimal control framework applied to robotic systems. Section 3 introduces a computational method for computing evolution of the state probability density function (PDF) of our model. In this section we compare the obtained results with the solution previously obtained from the PDE system describing the same evolution. Section 4 describes a computational method for evaluation of adjoint state distributions. Section 5 considers a special case of the process self-adjoint systems. Conclusion and suggestions for future work are provided in the last section.

2 Modeling and Control Framework

In the modeling framework we consider, the state of an individual robot at time t is uniquely defined by the couple $(x(t), q(t))$, $x \in X$, $X \in \mathbb{R}^n$, $q \in Q, Q = \{1, 2, ...N\}$. While in the discrete state (mode) $k \in Q$ the continuous state of robot obeys differential equation $\dot{x} = f_k(x, t)$. We also assume that switching among discrete states, say from the state $k \in Q$ to the state $j \in Q$, $(k \neq j)$ is described by stochastic transition rates λ_{kj}, and that $x(t)$ is a continuous function of time. In other words, the continuous state just before the discrete state transition $x(t^-)$ is equal to the state $x(t^+)$ after the state transition. This very general model of individual robot is illustrated in Fig. 1 and the modeling framework we are applying here is detailed in [10].

Recognizing that the state of the individual robot is composed of discrete and continuous components, the state probability density function (PDF) is a vector of functions $\rho(x, t) = [\rho_1(x, t)\ \rho_2(x, t) \ldots \rho_N(x, t)]'$. Each component $\rho_i(x, t)$ corresponds to the discrete state i, and the symbol (') denotes vector

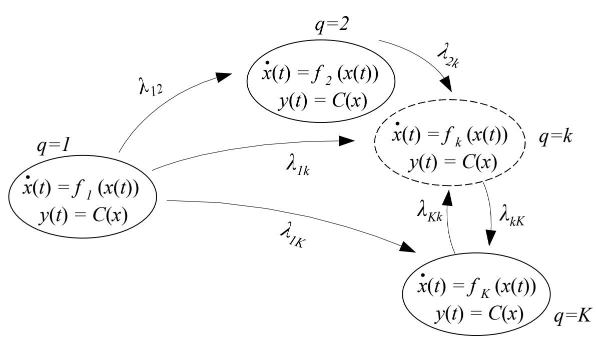


Figure 1: Stochastic hybrid automaton model of a robot in the probabilistic framework: discrete state q; continuous vector x vector field f_k, $k \in Q$ describes change of continuous state ; stochastic transition rates λ_{kj}, $k, j \in Q$ describe mode switching; y is the measurable output, if the full continuous state of the robot is measurable, C is unity matrix.

transpose. The state PDF satisfies

$$\sum_{i \in Q} \int_X \rho_i(x,t)dx = \sum_{i \in Q} P_i(t) = 1, \quad \text{where} \ \ P_i(t) = \int_X \rho_i(x,t)dx \tag{1}$$

where $P_i(t)$ is the probability of the discrete state i at the time point t. Let us define the vector of discrete probabilities $P(t) = [P_1(t), P_2(t), \dots P_N(t)]'$, then evolution of the probability vector is given by:

$$\dot{P}(t) = F_t(t)P(t), \quad \text{where} \ \ [F_t]_{ij} = \lambda_{ij}(t) \tag{2}$$

with matrix F_t defining the transition rates among the discrete states. In general, correspondence between the matrix F_t members $[F_t]_{ij}$ and the transition rates λ_{ij} is not one-to-one. Assuming that the transition rates depend on a vector $u(t) = [u_1(t) \ u_2(t) \ \dots u_M(t)]'$ of variables $u_i, i = 1, 2 \dots M$, we can define the transition rate matrix as a function of the vector $u(t)$, i.e., $F_t(t) = F_u(u(t))$. Consequently, the vector of the discrete state probabilities obeys [1]:

$$\dot{P}(t) = F_u(u)P(t) \tag{3}$$

Moreover, it can be proven [10] that the state PDF obeys the following system of partial differential equations (PDE):

$$\frac{\partial \rho(x,t)}{\partial t} = F(u)\rho(x,t) = (F_u(u(t)) + F_\partial)\rho(x,t) \tag{4}$$

where F_∂ is the diagonal linear differential operator. When the operator F_∂ is applied to $\rho(x,t)$, it results in:

$$[F_\partial \rho(x,t)]_{ij} = \begin{cases} \nabla \cdot (f_i \rho_i(x,t)), i = j \\ 0, i \neq j \end{cases} \qquad i, j = 1, 2 \dots N \tag{5}$$

Taking into account that the state PDF, $\rho(x,t)$, evolution depends on the vector $u(t)$ we can formulate the optimal control problem in the probability space using the cost function:

$$J = \int_X w'(x)\rho(x,T)dx \tag{6}$$

In this respect, the optimal control problem is the optimization problem:

$$u^*(t) = \max_{u(t)\in U_{ad}} J = \max_{u(t)\in U_{ad}} \int_X w(x)\rho(x,T)dx \tag{7}$$

Alternatively, to avoid the singular control problems [10] we can also consider the optimal control that includes the term penalizing the control:

$$u^*(t) = \max_{u(t)\in U_{ad}} J = \max_{u(t)\in U_{ad}} \int_X w(x)\rho(x,T)dx + \varepsilon \int_0^T u'(t)u'(t) \tag{8}$$

Anyway, the solution of this problem is a sequence of the optimal control $u^*(t)$, from the set of admissible control U_{ad}, such that the cost function is maximized. By suitable choice of the weighting function $w(x)$, the cost function can be used to find the optimal control maximizing probability of robot presence in the desired region of the robots' operating space.

The optimal control maximizing the criterion (6) is a special case of a more general optimal control problem of the evolution equation [4]. Under the condition that the operator $F(u)$ is bounded, i.e., $\|F(u(t))\| < \infty$ the minimum principle for PDEs can be applied [4]. According to the minimum principle, the optimal control $u^*(t)$ satisfies:

$$u^*(t) = \arg\min_{u\in U_{ad}} H(\rho(x,t), u(t), t) \tag{9}$$

In other words, for the optimal state PDF trajectory $\rho^*(x,t)$, the optimal control minimizes the Hamiltonian at each time point. The Hamiltonian is:

$$H(\rho(x,t),u,t) = \langle \pi(x,t), F(u)\rho(x,t)\rangle \tag{10}$$

where brackets $\langle \cdot, \cdot \rangle$ denote the scalar product of function vectors defined as:

$$\langle p(x), q(x)\rangle = \int_X p'(x)q(x)dx \tag{11}$$

The function vector $\pi(x,t)$ is the so-called adjoint state distribution and satisfies:

$$\frac{\partial \pi(x,t)}{\partial t} = -F'(u)\pi(x,t) \tag{12}$$

$$\pi(x,T) = -w(x) \tag{13}$$

The major limitation of the modeling approach presented here is that the system of PDEs, derived directly from the model shown in Fig. 1, is difficult to solve in general for any type of function $f_i(x)$ and of more than a single dimension. Nevertheless, the probabilistic approach we use is of fundamental importance for establishing mathematically tractable relation for the control of spatio-temporal distribution of large-size robotic populations. Once we establish this relation, we can introduce a different type of the propagator which is not necessarily in the form of differential equations. Ultimately, we can design an optimal control taking into account local interactions among the robots, or between the robots and the environment.

3 Stochastic Sampling Propagator

Evolution of the large-size population probability density function $\rho(x, t)$ is described by the PDE system (see Section 2, Eq.4). One way to obtain the evolution $\rho(x, t)$ is to solve the PDE system forward in time starting from an initial condition $\rho(x, 0) = \rho^0(x)$. This numerical solution is sensitive to function types used in the model, initial and boundary conditions as well as, in finite element approximation, to the mesh selection. Since numerical errors of solution accumulate, we also need to limit the terminal time T of the solution. If the numerical method for the solution is not carefully designed the influence of approximations and accumulated errors have almost random effect on the solution.

The Hamiltonian evaluation is instrumental for computing the optimal control based on minimum principle. In [9, 10] we formulated the algorithm for the optimal control in 1D case. However, it appears that sensitivity of numerical solutions contributing to the Hamiltonian evaluation is a limiting factor for solving the optimal control problem in more than one dimension and with larger number of discrete states.

Here we propose an approach to compute the evolution $\rho(x, t)$ based on stochastic trajectories of the hybrid state (x, q) evolution resulting from the model presented in Fig. 1, Section 2. To account for the fact that the transition rates can change in time, we assume that the control is a piecewise constant function of time discretized with the sample time ΔT. The basis for the proposed algorithm is the Gillespie's stochastic simulation algorithm [5].

To generate the trajectory of (x, q) we need to generate the initial state $(x(0), q(0))$ from the state PDF $\rho(x, 0) = \rho^0(x)$. Probability $P_i(t)$ of $q(t) = i$ is:

$$P_i(t) = \int_X \rho_i(x, t) dx \tag{14}$$

Therefore, the random variable $q(0) = i$ should be generated from the the discrete state probability distribution represented by the vector of discrete state probabilities $P(0) = [P_1(0)\ P_2(0)\ \dots P_N(0)]'$. Symbolically, we will represent it as:

$$q(0) = i \sim P(0) \tag{15}$$

Once the initial discrete $q(0)$ state is defined, the continuous variable $x(0)$ can be generated from the corresponding $\rho_i(x, 0)$ component of the state PDF ,i.e., from the probability $\mathcal{P}$ of $x(t)$ given that $q(t) = i$ and $t = 0$

$$x(0) \sim \mathcal{P}\{x|q(t) = i, t = 0\} = \rho_i(x, 0)/P_i(0) \tag{16}$$

Whenever the discrete state is $q(t) = i$, the evolution of the continuous state x obeys $\dot{x} = f_i(x)$. Therefore, generating trajectory $(x(t), q(t))$ reduces to the problem of generating the state transitions of the discrete state $q(t)$. Let us assume that at time $t = t_s$, $t_s \in [(k-1)\Delta T, k\Delta T)$ the hybrid state is $(x(t), q(t))$ the time instant at which the state changes t_c can be generated based on the following two rules:

(a) $t_c = t_s + t_t$, $t_t \sim e^{-t \sum_j \lambda_{ij}(k-1)}$, under condition that $t_c < k\Delta T$. If the condition is not satisfied, apply rule (b).

(b) $t_c = k\Delta T + t_t$, $t_t \sim e^{-t \sum_j \lambda_{ij}(k)}$, under condition that $t_c < (k+1)\Delta T$. If the condition is not satisfied increase k by 1. Apply rule (b) until the condition is satisfied.

These two rules define the time point t_c at which the jump from discrete state i to discrete state j happens, but do not specify the variable j. The state j needs to be sampled from the discrete state probability density function, i.e., from the probability $\mathcal{P}$ of $q(t^+) = j$ given that $q(t) = i$ provided in the vector of the discrete probability distribution with $N-1$ elements:

$$j \sim \mathcal{P}\{q(t)|q(t^-)=i\} = \underbrace{\left[\frac{\lambda_{i1}}{\sum_{n=1}^{N}\lambda_{in}}, \ \frac{\lambda_{i2}}{\sum_{n=1}^{N}, \ \lambda_{in}} \cdots \frac{\lambda_{iN}}{\sum_{n=1}^{N}\lambda_{in}}\right]}_{N-1} \tag{17}$$

The algorithm just described can be used to generate a single trajectory for the stochastic model shown in Fig. 1. In the limit of large number of samples, the normalized density of trajectory points will correspond to the solution the PDE system given by Eq.4 in Section 2. In this respect, the stochastic simulation is a computational propagator of the evolution $\rho(x,t)$ and we can denote it as:

$$\frac{\partial\rho}{\partial t} = F_{sim}(u(t))\rho \tag{18}$$

To illustrate and verify the algorithm for generating stochastic trajectories $(x(t), q(t))$ we use the following 1D and 2D examples.

3.1 1D example

The stochastic model presented in Fig. 2b illustrates the state PDF evolution of the large-size robot population along one dimension, Fig. 2a in which u_1, u_2 and u_3 correspond to stochastic rates of commands: move-left, move-right

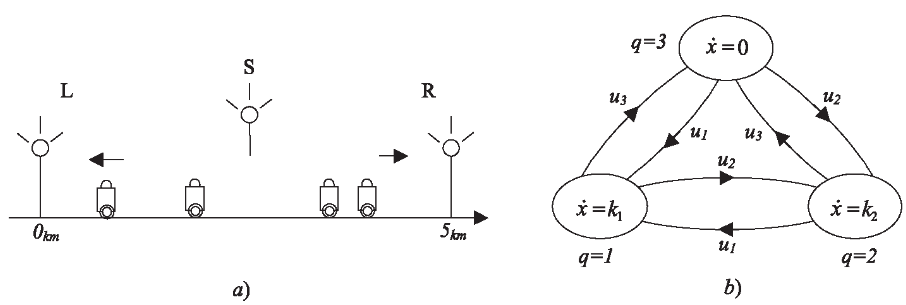


Figure 2: 1D example

and stop. In this example $k_1 = -0.5$ and $k_2 = 0.25$. The control $u(t) = [u_1(t)\ u_2(t)\ u_3(t)]$ is computed as the optimal control based on the Minimum-principle and Hamiltonian presented in the previous section.

The cost function is:

$$J = \int_X w'(x)\rho(x,t)dx + \varepsilon \int_0^T u_1^2(t) + u_2^2(t) + u_3^2(t)dt \tag{19}$$

where $\varepsilon = 10^{-7}$, the weighting $w(x) = [0\ 0\ w_3(x)]'$ and initial condition $\rho(x,0) = [0\ \ 0\ \rho_3(x,0)]'$ are defined by:

$$w_3(x) = \begin{cases} \frac{1}{\sqrt{0.01}} exp(-\frac{(x-1.75)^2}{0.01}), & 1.25 < x < 2.25 \\ 0, \ otherwise \end{cases} \tag{20}$$

$$\rho_3(x,0) = \begin{cases} \frac{1}{\sqrt{0.02\pi}} exp(-\frac{(x-2.5)^2}{0.02}), & 2 < x < 3 \\ 0, \ otherwise \end{cases} \tag{21}$$

The optimal control sequence $u^*(t) = [u_1^*(t)\ u_2^*(t)\ u_3^*(t)]$ in the time interval $0 < t < 3$ is defined by:

$$u_1^*(t) = \begin{cases} 2, \ t_1 < t < t_2 \\ 0, elsewhere \end{cases}, \quad u_2^*(t) = 0, \quad u_3^*(t) = \begin{cases} 2, \ t_1 < t < t_2 \\ 0, elsewhere \end{cases} \tag{22}$$

The evolution of the state PDF for this system under the control $u^*(t)$ is presented in Fig. 3. We present only $\rho_1(x,t)$ and $\rho_3(x,t)$ because under this control $\rho_2(x,t) = 0,\ \forall t$.

For the illustration we generated 10 stochastic trajectories of continuous variable x (see Fig. 4) under the control $u^*(t)$. Evolution of the discrete state q can be observed from the trend in x. When x decreases the discrete state is 1, and when it stays constant the state is $q = 3$. It is worth mentioning that

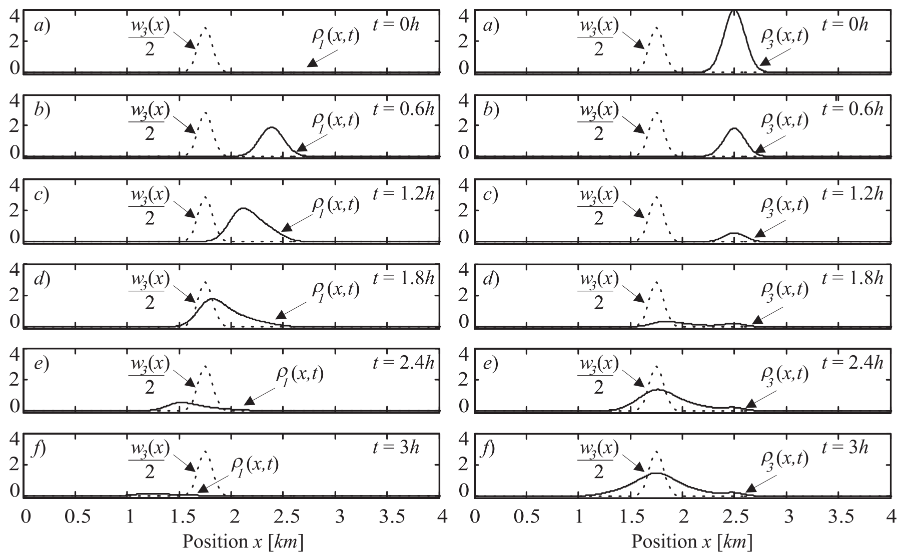


Figure 3: The finite element solution of the state PDF evolution for 1D example under the optimal control $u^*(t)$

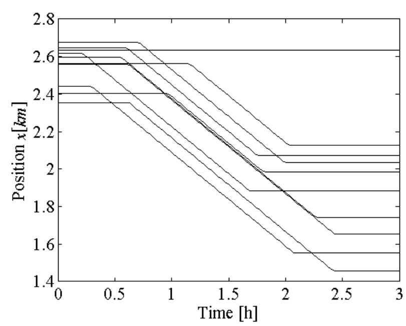


Figure 4: Random set of 10 trajectories resulting from the stochastic simulation under the optimal control $u^*(t)$

among these 10 trajectories there is one for which $x(t)$ is constant. The small pick at around the point 2.5 in the right panel of the Fig.3 at $t = 3$, confirms that the probability of such trajectories is non-zero, but it is small.

To obtain the state PDF $\rho(x,t)$, i.e., its components $\rho_i(x,t)$ at specific time point t we need to collect points $x(t)$ and estimate components $\rho_i(x,t)$. It is obvious that 10 trajectories cannot provide a good estimate of $\rho(x,t)$ For this reason, we generated 10^5 trajectories and computed histogram probability density function estimate. That means that we discretized the x axis into intervals of the length Δx and counted how many points fell into a specific region. Finally, we normalized the histogram so that the estimated $\rho(x,t)$ satisfied condition 1, given in Section 2. Results are presented in Fig. 5. As expected the match between the numerical PDE system solution and the result obtained from stochastic trajectories is exact. There are only negligible discrepancies due to data sampling from finite number of trajectories.

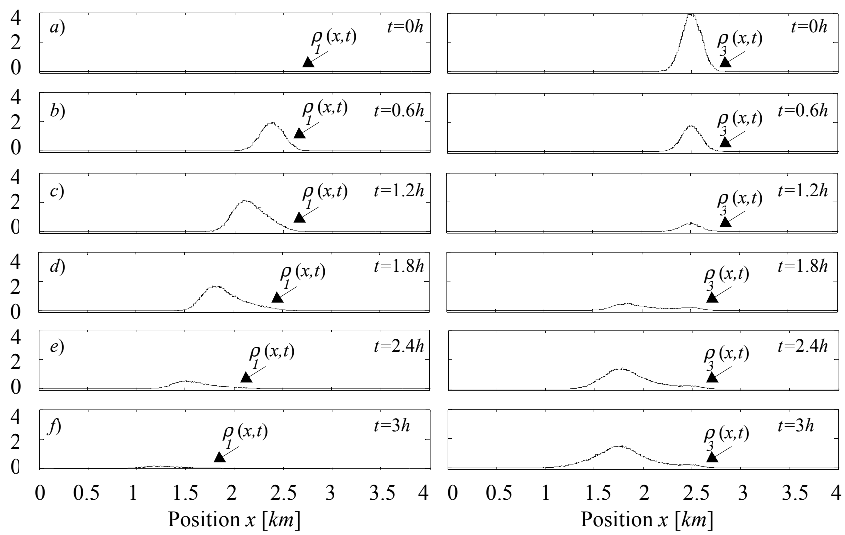


Figure 5: The stochastic simulation solution of the state PDF evolution for 1D example under the optimal control $u^*(t)$

3.2 2D example

The benefit of applying stochastic sampling algorithm and the reason for introducing the propagator is its ability to propagate multi-dimensional probability density function. To illustrate that, we apply stochastic sampling for computing distribution of the model presented in Fig. 6. The model is based on a scenario in which a population of small robots is initially concentrated on a location in the operating region.

The robots are controlled by the stochastic signals produced by aerial robots (Fig. 6*a*). Each robot in the population moves in the direction of the active signal source. Under the assumptions that aerial robots are far away from the population and that the robot velocity is constant ($v = 1$), the robotic motion model is given by the stochastic model shown in Fig. 6*c*. In this example no direct transitions between states 1 and 3 exist. The model and scenario are detailed in [10, 11]

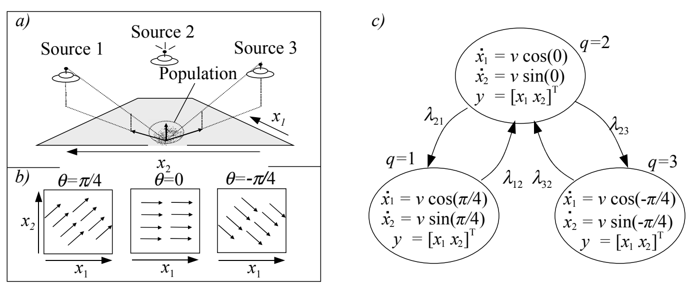


Figure 6: 2D example

We take the case in which $\lambda_{12} = 0.5$, $\lambda_{21} = 0.1$, $\lambda_{23} = 0.9$ and $\lambda_{32} = 0.1$ are constant. The trajectories are generated based on the proposed algorithm and sampled at time points t=0, 0.39, 0.79, 1.18, 1.57, 1.96. The samples are presented in Fig. 7 and the color of dots denotes the time point of the sample. To produce this result, we used only 100 samples so that difference in the density of points could be observed.

The transition rates of this example are set without consideration of optimal control. Ultimately we should be able to control these 2D distributions similar to the ones in the 1D example presented in the previous section. For comparison, we also present in Fig. 8 the state PDF evolution resulting from the solution of the corresponding PDE system. This evolution can help us recognize possible

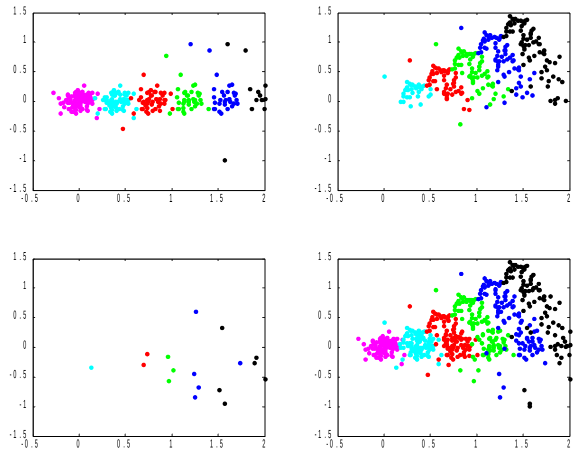


Figure 7: The stochastic simulation solution of the state PDF evolution for 2D example (100 points). The transition rates are $\lambda_{12} = 0.5$, $\lambda_{21} = 0.1$, $\lambda_{23} = 0.9$ and $\lambda_{32} = 0.1$. The time points are t=0, 0.39, 0.79, 1.18, 1.57, 1.96 and are represented by different colors from pink to black; dimensions along x and y axis are in km.

but improbable stochastic trajectory realizations presented in Fig. 7. We can also note the polygon like contour plots resulting from the imprecision of the numerical solution of the PDE system.

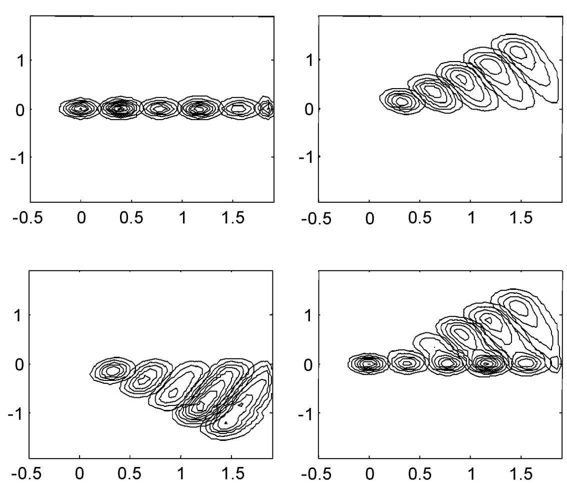


Figure 8: Numerical PDE system solution

4 Stochastic Sampling for Adjoint State Distribution

The Hamiltonian, corresponding to the optimal control problem Eq.7, Section 2, is defined by:

$$H(t) = \int_X \pi(x,t)F(u)\rho(x,t)dx = E_\rho\left\{\pi(x,t)F(u)\right\} \tag{23}$$

and can be seen as an expected value of the adjoint state (or co-state) with respect to the the state PDF. Assuming that the adjoint state evolution $\pi(x,t)$ is known, evaluation of the expected value is easy having in mind that we know how to generate realizations of the hybrid state stochastic trajectory $(x(t), q(t))$. One way to obtain the adjoint state distribution evolution $\pi(x,t)$ is to solve the PDE system Eq. 12, Section 2, backward in time starting from the initial condition $\pi(x,T) = -w(x)$. However, we should notice that the complexity of solving the adjont PDE system is as complex as solving the state PDF evolution $\rho(x,t)$. Therefore, it is natural to consider an opportunity to solve adjoint state evolution $\pi(x,t)$ using stochastic sampling similar to the one presented in the previous section.

The adjoint set of equation is defined by:

$$\frac{\partial \pi(x,t)}{\partial t} = -F'(u)\pi(x,t) \tag{24}$$

$$\pi(x,T) = -w(x) \tag{25}$$

Let us assume that the $\pi(x,t)$ as well as $w(x)$ are from the state PDF class of functions, i.e., they satisfy the following property (*Property 1*):

$$\sum_{i\in Q}\int_X \pi(x,t)dx = \sum_{i\in Q} P_i^\pi(t) = 1 \tag{26}$$

$$\sum_{i\in Q}\int_X w(x)dx = \sum_{i\in Q} P_i^\pi(T) = 1 \tag{27}$$

where $P_i^\pi(t)$ is the probability of discrete state i at time t assuming that the $\pi(x,t)$ is the state PDF of appropriately defined stochastic process. If the assumption is valid than the discrete state probability evolution backward in time is defined as:

$$\dot{P}^\pi(T-\tau) = F_u'(u)P^\pi(T-\tau) \tag{28}$$

but here we find the contradiction. The transpose, F_u', of the transition rate matrix F_u does not warranty in general *Property 1.* Therefore, the adjoint PDE system does not define a conservation law in the probability space. This is a major obstacle for the straight-forward design of an algorithm that will evaluate adjoint state distribution evolution based on the stochastic process similar to the one for evaluating the state PDF. Consequently, an alternative approach needs to be considered.

The algorithm for computing solution of adjoint equation has been recently published in [14]. The algorithm is based on an assumption that the PDE system describing the adjoint state distribution evolution

$$\frac{\partial \pi}{\partial t} = -F'(u)\pi(x,t) \tag{29}$$

can be discretized in space and time, so that we obtain a system of linear equations defining the solution of Eq. 29:

$$\bar{A}\bar{\pi} = \bar{b} \tag{30}$$

where $\bar{\pi}$ is a vector of values $\pi_i(m\Delta X, k\Delta T)$, $i \in Q$, ΔT is the discretization sample time and m corresponds to the mth element of the space discretization. Obviously the matrix $\bar{A}$ is a square matrix of dimension $(|M|\cdot|Q|\cdot K)\times(|M|\cdot$

$|Q| \cdot K)$, where $|M|$, $|K|$ and $|Q|$ are the number of discrete element of space X, the number of the time steps and the number of the discrete states, respectively. Vector $\bar{b}$ is equal to 0 except for the elements corresponding to the time T where $\bar{b}$ element values correspond to $-w(x)$.

Equation 29 is an algebraic equivalent of Eq. 30. Based on this algebraic interpretation, $\bar{\pi}$ can be obtained as:

$$\bar{\pi} = \bar{A}^{-1}\bar{b} \tag{31}$$

Instead of computing the inversion algebraically, the authors of [14] proposed an algorithm in which Eq. 30 is rewritten in the following form:

$$\bar{\pi} = \bar{A}\bar{\pi} + b, \quad A = I - P\bar{A}, \text{ and } b = P\bar{b} \tag{32}$$

with P a being block-diagonal preconditioner matrix. Based on this form the solution for $\bar{\pi}$ can be expanded in a Neumann series:

$$\bar{\pi} = b + Ab + A^2b + A^3b + \ldots \tag{33}$$

This series converges only if the spectral radius of A is less than 1. In this case we can use the Monte Carlo method to sample this infinite series by a Markovian random walk.

The algorithm we present here is adopted from [14] and the specific values are adjusted to the special case of dynamical system we are dealing with. The single realization on the random walk we denote by $[p]$. The value $W[p]$, or $D[p]$ denotes the variable W, or D related to the single realization of the random walk. Variable $\alpha_k[p]$ denotes the state of the random walk at the step k from the realization of the random walk $[p]$. Introducing this notation, the algorithm to compute $\bar{\pi}$ is as follows:

1. Initialize $W[p] = 1$, $D[p] = 0$ and set $t = 1$

2. Compute the matrix A, and based on it, define probabilities of state transitions:
$$p(i,j) = \frac{|A_{ij}|}{\sum_j A_{ij}} \tag{34}$$

3. For each random walk $[p]$ that is at time t choose the next state based on random probabilities $p(i,j)$. If the state $\alpha_{k+1}[p]$ is not the final exit state, update $W[p]$ and $D[p]$ as follows:
$$W[p] = W[p] w_{\alpha_k[p]\alpha_{k+1}[p]}; \quad D[p] = D[p] + W[p] b_{\alpha_{k+1}[p]} \tag{35}$$
If $\alpha_{k+1}[p]$ is the final exit state, the random walk is absorbed and we freeze $D[p]$.

4. Repeat state 3 until all random walks at time step t are either absorbed or left the time step. If $t < m$, then let $t = t + 1$ and go to step 2.

5. After being done with the last time step m, all random walks are absorbed. Compute the sample mean of the estimator $\frac{1}{q}\sum_{p=1}^{q} D[p]$, which is approximation of $\bar{\pi}$.

Evaluation of the state PDF as well as adjoint state distribution based on stochastic simulation provides an opportunity for the time efficient computing of the optimal control. However, computations need to be carefully analyzed since the adjoint state evolution is not a conservation law and it can result in computational instability. Hopefully, a sacrifice of precision of the adjoint state distribution computations would not influence much the Hamiltonian evaluation. Based on our experience, the state PDF evaluation presented in the previous section is much faster than the evaluation based on solving partial differential equations. Similar tests should be done with the computation of the adjoint state distributions presented in this section. With the design of dedicated hardware, we believe that all this difficulty can be avoided and ultimately complete stochastic sampling decision making can be done on-board and in real-time.

5 Process Self-adjoint Systems

In the previous section we explained that in general the adjoint state equation is not a conservation law. Consequently the stochastic sampling computation of adjoint state distributions is a more difficult job than estimation of the state PDFs. In this section we consider a special type of dynamical systems in which the adjoint state evolution is a conservation law, and we call this type of systems process self-adjoint.

Let us assume that that the state pdf evolution is described by:

$$\frac{\partial \rho(x,t)}{\partial t} = F'(u)\rho(x,t) \tag{36}$$

$$\rho(x,0) = \rho^0(x) \tag{37}$$

and the adjoint state distribution evolution is given by:

$$\frac{\partial \pi(x,t)}{\partial t} = -F'(u)\pi(x,t) \tag{38}$$

$$\pi(x,T) = -w(x) \tag{39}$$

where the operator $F(u) = (F_u(u(t)) + F_\partial)$ and the operator $F'(u)$ is:

$$F'(u) = (F_u(u(t)) + F_\partial)' = (F'_u(u(t)) - F_\partial) \tag{40}$$

In other words, the adjoint state distribution evolution is rewritten as:

$$\frac{\partial \pi(x,t)}{\partial t} = -(F_u'(u(t)) - F_\partial)\pi(x,t) \quad (41)$$

$$\pi(x,T) = -w(x) \quad (42)$$

If we rewrite this equation backward in time we obtain:

$$\frac{\partial \pi(x,T-\tau)}{\partial \tau} = (F_u'(u(t)) - F_\partial)\pi(x,T-\tau) \quad (43)$$

$$\pi(x,\tau=0) = -w(x) \quad (44)$$

Finally, introducing the substitution $\phi(\tau) = -\pi(x, T-\tau)$ we can obtain the backward solution of the adjoint state evolution in the form of a forward time PDE system:

$$\frac{\partial \phi(x,\tau)}{\partial \tau} = (F_u'(u(t)) - F_\partial)\phi(x,\tau) \quad (45)$$

$$\phi(x,0) = w(x) \quad (46)$$

Let us consider the special case in which the discrete state transition matrix satisfies $F_u'(u(t)) = F_u(u(t))$. Then we can conclude that the adjoint state evolution Eq. 5 (i.e., its equivalent backward time evolution Eq. 5) is conservation law. Consequently, if $w(x)$ is a function of the same type as the state PDF, then the adjoint state distribution π, i.e., its equivalent ϕ can be considered the co-state probability density function of some stochastic process which we will explain briefly.

The condition $F_u'(u(t)) = F_u(u(t))$ implies that the transition rate from the discrete state i to the discrete state j is the same in the opposite direction, i.e., from the state j to the state i. Interestingly enough, under assumption that

$F'_u(u(t)) = F_u(u(t))$ the robotic system can be considered process self-adjoint. In other words, the adjoint state evolution corresponds to a stochastic process. For example, let us consider the vehicle of which dynamics can be represented by three discrete and three continuous states (see Fig. 9a and vehicle illustration in Fig. 10). The first discrete state corresponds to the vehicle moving without

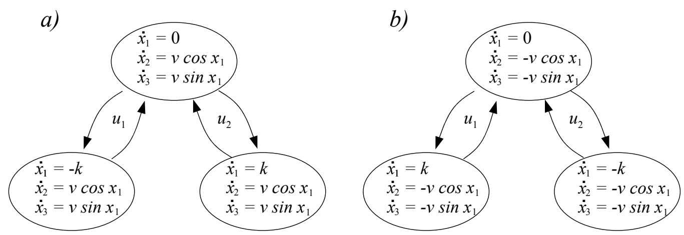


Figure 9: a) The stochastic process describing the evolution of the state PDF, b) The stochastic process describing the evolution of the adjoint state PDF backward in time ϕ

change of direction. The second and the third state correspond to the increase and decrease of the angle of the moving direction, respectively. To obtain a stochastic process describing evolution of the adjoint state distributions we need each dynamics of the discrete state i, f_i substitute with $-f_i$ and keep in mind that this stochastic process describes the adjoint state backward in time. The resulting adjoint process is depicted in Fig. 9b. The generation of this stochastic process can be performed using the algorithm presented in Section 3.

To summarize, evaluation of the state $\rho(x,t)$ and co-state $\pi(x,t)$ PDFs for this case will be based on stochastic samples from trajectories defined by models presented in Fig. 9. The trajectories need to be generated and collected in 3D since we deal with the three continuous state variables. The 2D projection of these 3D trajectories is illustrated in Fig. 10. The distribution of initial points of trajectories generated forward in time relates to $\rho^0(x)$ in the region A (see Fig. 10) while the initial distribution of trajectories generated backward in

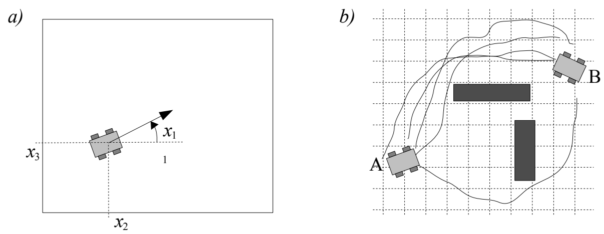


Figure 10: a) The vehicle and continuous state variables PDF, b) The forward and backward paths starting from the region A and B, respectively

time relates to $w(x)$ in the region B. In this respect the Hamiltonian Eq. 23 measures the match between forward and backward trajectories and its value can be used for control improvement, i.e., to provide a better match. Although we discuss the special case, this is an elegant example providing rationale behind the ant-optimization [3] algorithms.

6 Conclusions and Future work

In this research project we study a theoretical possibility of solving the optimal control problems based on stochastic sampling. We came to the conclusion that the main difficulty in this approach is evaluation of the adjoint state distributions. While proposed Monte Carlo algorithm is numerically efficient, its convergence is not guaranteed. An alternative approach to avoid the difficulty is to consider a class of robotic systems for which the adjoint state distribution evolution corresponds to the stochastic process. We call such systems process self-adjoint. In this case both state PDF and co-state PDF contributing to the Hamiltonian can be evaluated based on stochastic samples. Based on the process self-adjoint systems we can establish a relation between the stochastic sampling stochastic optimal control and ant-optimization algorithms [3].

Future work along this approach will be focused on examples based on algorithms presented here. We are also considering establishing relation between the stochastic sampling methods and Hamilton-Jacobi-Bellman equation in order to solve the stochastic optimal control problems in a closed loop.

The main idea behind our work is to solve stochastic optimal control problems based on stochastic sampling. We believe that in this way we can employ modern statistical methods and control theory to solve complex stochastic optimal control problems for multi-robot systems in real-time.

References

[1] C. Cassandras and S. Lafortune. *Introduction to Discrete Event Systems.* Kluwer Academic Publ., 1999.

[2] A.J. Van der Schaft and J.M. Schumacher. *An Introduction to Hybrid Dynamical Systems.* Springer-Verlag, 2000.

[3] M. Dorigo and T. Stützle. *Ant Colony Optimization.* MIT Press, 2004.

[4] H.O. Fattorini. *Infinite Dimensional Optimization and Control Theory.* Cambridge University Press, 1999.

[5] D. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[6] Y.C. Ho and A. E. Bryson. *Applied Optimal Control: Optimization Estimation and Control.* John Wiley & Sons, 1975.

[7] H. J. Kappen. Linear theory for control of nonlinear stochastic systems. *Physical Review Letters*, 95(20):200201, 2005.

[8] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time.* Springer, 2001.

[9] D. Milutinović and P. Lima. Modeling and optimal centralized control of a large-size robotic population. *IEEE Transactions on Robotics*, 22:1280–1285, 2006.

[10] D. Milutinović and P. Lima. *Cells and Robots : Modeling and Control of Large-Size Agent Populations*. Springer, 2007.

[11] D. Milutinović, P Lima, and M. Athans. Biologically inspired stochastic hybrid control of multi-robot systems. In *Proceedings of the 11th International Conference on Advanced Robotics*, 2003.

[12] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.

[13] B. van den Broek, W. Wiegerinck, and B. Kappen. Graphical model inference in optimal control of stochastic multi-agent systems. *Journal of Artificial Intelligence Research*, 32:95–122, 2008.

[14] Q. Wanga, D. Gleicha, A. Saberia, N. Etemadib, and P. Moin. A monte carlo method for solving unsteady adjoint equations. *Journal of Computational Physics*, 227(12):6184–6205, 2008.

Appendix A

Title : Kalman Smoother Based Force Localization and Mapping using Intravital Video Microscopy

Authors : Dejan Lj. Milutinović and Devendra P. Garg, Department of Mechanical Engineering and Materials Science Duke University, Durham, NC, USA

Submitted for publication to: Journal of Dynamic Systems, Measurement, and Control, Transactions of the ASME, Series G, 2009.

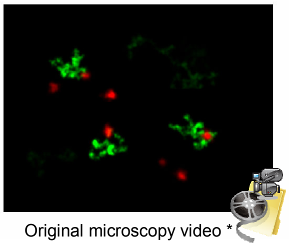
Abstract

Motility is an important property of immune system cells. It provides cells with the ability to perform their function not only at the right time, but also at the right place. In this paper, we introduce the problem of modeling and estimating an effective force field directing cell movement by the analysis of intravital video microscopy. A computational approach is proposed for solving this problem without dealing with a parameterized spatial model of the field in order to avoid potential errors due to inaccurate spatial model assumptions. We consider the dynamics of cells similar to the dynamics of distributed agents typically used in the field of swarm robotics. The method utilizes a fixed-interval Kalman filter based smoother. Its application results in a map giving the intensity and direction of the effective force field. The results show that real-time video images are a source of data, enabling us to visualize intriguing spatio-temporal phenomena inside immune system organs. The proposed approach can fill the existing gap between contemporary technology and quantitative data analyses present in the field of biosystems.

Appendix B

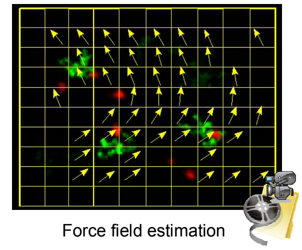
FORCE LOCALIZATION AND MAPPING FROM INTRAVITAL VIDEO MICROSCOPY

Dejan Milutinović, Devendra P. Garg
Robotics and Manufacturing Automation Laboratory, Duke University



Original microscopy video *

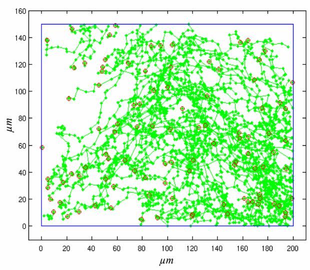
- **Video microscopy** provides information about immune system cells residing in tissues of living organisms.
- **F**orce **L**ocalization **A**nd **M**apping (FLAM) is the **problem of estimating an effective force field influencing cell motility** from intravital video microscopy.
- For solving the FLAM, we propose a **computational approach** without dealing with a parameterized spatial model of the field. In this way, we avoid potential errors due to inaccurate spatial model assumptions.
- The **estimation is based on the individual agent motility model**.
- **No alternative microscopy method exists** for measuring these forces.
- The **computational method** does not interfere with the physiological condition inside organs.



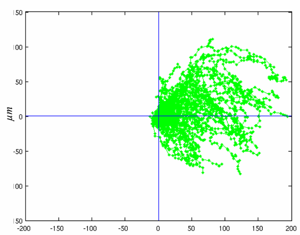
Force field estimation

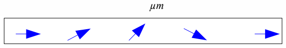
* Figure from D. Milutinović, P. Lima, “Cells and Robots”, Springer, 2007

Data



Analysis

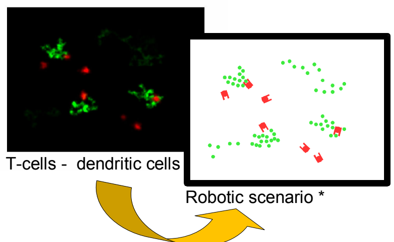




Force profile:
horizontal component constant
vertical component sinusoidal

Standard approach:
data aligned to the point (0,0)
analysis of the obtained track distribution

In this example, we can easily identify the horizontal force component, but the analysis does not provide an insight into the vertical force component.



T-cells - dendritic cells

Robotic scenario *

Agent Dynamics Model

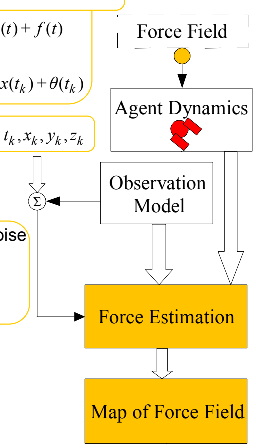
Position : $\dot{x} = v(t)$

Velocity : $\dot{v} = -cv(t) + \sigma\xi(t) + f(t)$

Force : $\dot{f} = \sigma_F \xi_F(t)$

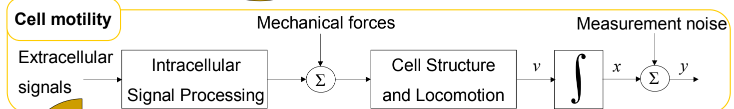
Measured positions: $y(t_k) = x(t_k) + \theta(t_k)$

Data : t_k, x_k, y_k, z_k

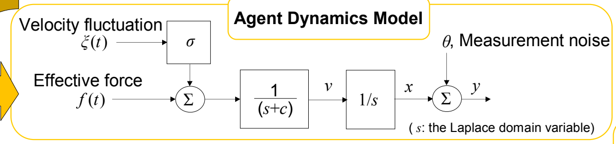


Data are known; this is a smoothing, not a filtering problem

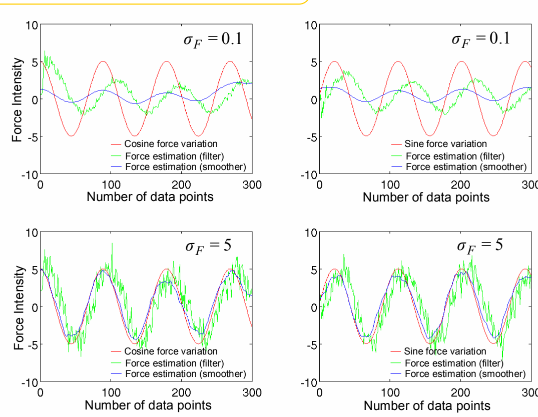
Cell motility



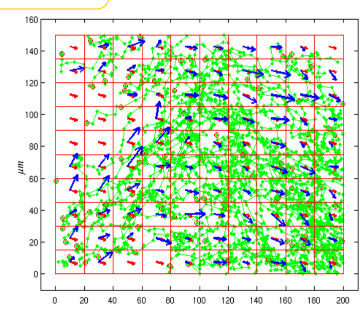
Agent Dynamics Model



Single cell track estimation performance



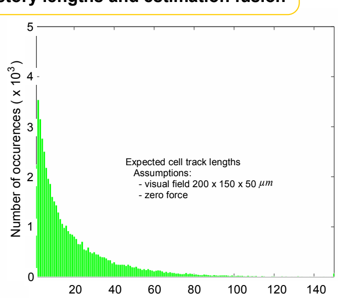
Results



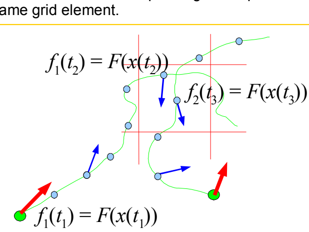
Algorithm

1. Estimate the force as if a constant for each trajectory i (Smoother $\sigma_F = 0$)
2. “Average” all estimates to obtain the constant force term
3. Estimate the force along each i[th] trajectory initially assuming that it has value of the constant force obtained in Step 2 above (Smoother $\sigma_F \neq 0$)
4. “Average” all estimates that belong to the same grid cell

Trajectory lengths and estimation fusion



The cell tracks are short in comparison to the number of data points necessary for reliable force estimation; We fuse estimations corresponding to the positions of the same grid element.



$$f_1(t_2) = F(x(t_2)) \approx F(x(t_3)) = f_2(t_3)$$

Conclusions

- An optimal smoother applied to force estimations as sensed by individual cells.
- Estimations averaged over a grid of rectangular regions taking into account the estimation covariances.
- The force field is estimated and visualized by the average forces assigned to each rectangular region. No analytical constraints on the force field.
- To apply the method to the analysis of imaging data routinely, a rapid adjustment of the estimator parameters is important.
- Our results show that real-time video images are a source of data, enabling us force visualization, which can be used for studying intriguing spatio-temporal phenomena inside immune system organs.
- The problem formulation and the solution fill the existing gap between contemporary technology and quantitative data analyses present in the field.

The CMPI Symposium on Multi-Scale Modeling of Host/Pathogen Interactions, June 23 - 25, 2009, University Club, Pittsburgh, PA
Acknowledgment: This research was supported by the Army Research Office under Short Term Innovative Research (STIR) grant number W911NF-07-R0003.